ELSEVIER

Journal of Chromatography B, 725 (1999) 17–37

JOURNAL OF CHROMATOGRAPHY B

Review

Physico-chemical and biological analysis of true combinatorial libraries

J.A. Boutin[a,*], P.H. Lambert[b], S. Bertin[b], J.P. Volland[b], J.L. Fauchère[c]

[a]*Department of Molecular and Cellular Pharmacology, Institut de Recherches SERVIER, 125 Chemin de Ronde, 78290 Croissy-sur-Seine, France*

[b] *Department of Analytical and Physical Chemistry, Institut de Recherches SERVIER, 11 rue des Moulineaux, 92150 Suresnes, France*

[c]*Department of Peptide and Combinatorial Chemistry, Institut de Recherches SERVIER, 11 rue des Moulineaux, 92150 Suresnes, France*

Abstract

Combinatorial libraries offer new sources of compounds for the research of pharmacological agents such as receptor ligands, enzyme inhibitors or substrates and antibody-binding epitopes. The present review stresses the main roles played by both physico-chemical analysis, particularly when complex mixture of compounds are synthesized as libraries, and biological analysis from which active compounds are identified. After a brief discussion of semantic problems related to the designation of the product mixtures, the physico-chemical analysis of mixtures is reviewed with special emphasis on mass spectrometric techniques. These methods are able both to give a representative view of a library composition and to identify single critical compounds in large libraries. Then the biological screening of such combinatorial libraries is critically discussed with respect to the power and limitations of the methods used for the identification of the active components. Special attention is given to the complex process of library deconvolution. It is pointed out that while combinatorial techniques have evolved towards sophisticated high-tech methods, simple and robust biochemical tests should be used to deconvolute. From a large panel of published examples, a set of trends are identified which should help investigators to choose the most appropriate assay for the discovery of new entities.

Keywords: Reviews; Combinatorial libraries; Library deconvolution

Contents

*Corresponding author. Tel.: +33-1-41182748; fax: +33-1-41182740

0378-4347/99/$ – see front matter
PII: S0378-4347(98)00569-6

1. Introduction

Contemporary medicinal chemistry greatly benefits from the methods recently developed for the multiple simultaneous synthesis of large numbers of compounds from which hits, leads and finally candidate drugs are identified. At each and every step of drug discovery, a special niche for chemical analysis is identified which plays a key role in the drug discovery process (Fig. 1). The wide variety of synthetic methods generating a large molecular diversity in libraries of compounds is the core of combinatorial chemistry.

A word of warning is necessary, in the first place, about the nomenclature used in the recent literature. The term combinatorial is not used in a consistent way and refers often both to the production of collections of individual compounds and to that of compound libraries, i.e. of mixtures of structurally related compounds [1–4]. Strictly speaking, the word combinatorial should be confined to the description of synthesis procedures in which at least one step produces one or several mixtures (generally called libraries) of products by some randomizing process as this is typically the case in split synthesis described first by Furka et al. [5]. In contrast, parallel synthesis, although able to produce large number of compounds, does not lead to mixtures but to collections often called arrays of products. The present short review deals with combinatorial libraries, such as those laid out by Furka et al. [5] and Houghten et al. [6]. When our group came into the field of combinatorial peptide libraries, we intended to synthesize relatively small peptide libraries of tetra- to hexapeptides, using a robotic instrument. To our surprise, the characterization of the libraries described in the literature was not an important topic. Nevertheless, we tentatively analyzed our libraries by several methods [7] and we are still considering today that chemical analysis is a key point for the characterization of complex mixtures. The reasons for the analytical treatment of libraries being underregarded may lie both in the optimized conditions of peptide synthesis of which the success is taken for granted, and also in the complex analytical data obtained with product mixtures from which specific information about the product distribution is difficult to extract. This has even led to the statement that ‘mixture analysis methods are of little value for interesting sized libraries’ [8]. By contrast, we tend to show in this review that physico-chemical analysis using modern and upcoming techniques is essential to characterize interesting size libraries and to check the matching of their expected to their experimentally observed composition.

Another point of interest, surprisingly poorly considered in the literature, is the comparative power of the biological screening tools used for the identification of the active components in complex combinatorial libraries. On the basis of a number of

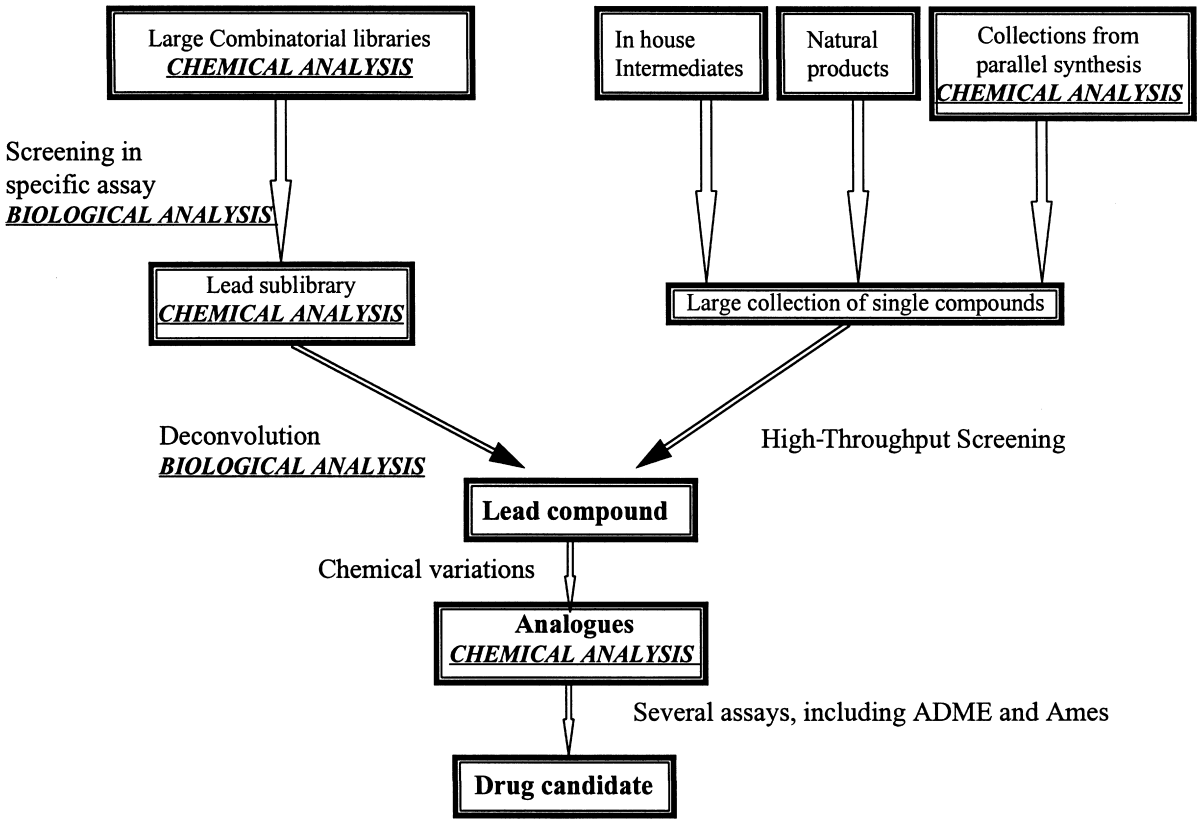

Fig. 1. Schematic representation of the various steps of modern drug discovery.

reported experimental studies, we tend to establish a hierarchy of the biological testing systems as to their specificity and relevance for ligand identification in complex mixtures.

2. Physico-chemical analysis of combinatorial libraries

Since the pioneering work of Geysen et al. [9], the synthesis and screening of peptide mixtures has become a common procedure in many industrial and academic laboratories. The first combinatorial syntheses of peptide mixtures using a mix and divide strategy (split synthesis) have been performed by Furka et al. [5], Houghten et al. [6] and Lam et al. [10], and further developed after 1991 by many others (for a review, see Fauchère et al. [11]). The initial enthusiasm for the possibility of screening peptide mixtures using iterative deconvolution methods corresponded to a blossom of publications. A large number of papers described the discovery of new biologically active ligands of enzymatic or receptor macromolecules. Some of the physical and statistical limitations related to the size of the synthetic libraries were also investigated by Zhao et al. [12] and by Boutin and Fauchère [13]. In the following section, we give an overview of the analytical characterization of compound mixtures such as those obtained by combinatorial methods, with an emphasis on polyamide (peptide) libraries.

For the identification of the most active component(s) by deconvolution to be reliable, all expected compounds must be present in equimolar amounts and free of side-products. Although peptide libraries prepared by the the well-controlled Merrifield solid-phase synthesis [14], some failure sequences due to incomplete deprotection or incomplete coupling or to other side reactions can occur. However, as can be seen from Table 1 (column

Table 1
Representative compendium of biologically tested combinatorial libraries

Library type[a]	Alphabet[b]	Complexity[c]	Target[d]	Assay type[e]	Deconvolution[f]	Hit[g,h]	Analysis[i]	References
1/PEPTIDES								
Hexapeptide	19	34 012 224	Monoclonal antibody	ELISA	iSURF	Ac-DVPDYA-NH2	None reported	Houghten et al., 1991 [6]
Decapeptide	20	4 000 000 000 000	Anti-haemagglutinin Ab	ELISA	Positional scanning	Ac-DDDDDVPDYA-NH2	AAA	Pinilla et al., 1994 [15]
Tetrapeptide	19	130 321	anti β-endorphin	ELISA	Sequencing after orthogonal release	YGGFG, YGVFG	None	Salmon et al., 1993 [16]
Tetrapeptide	19	6859	gpIIb/IIIa	ELISA	Sequencing after orthogonal release	CRGDC, GARYC	None	Salmon et al., 1993 [16]
Pentapeptide	26	11 881 376	IgAC5	ELISA	Positional scanning	HFVQH	None reported	Bianchi et al., 1995 [17]
Octapeptide	19	14 000 000	RSV epitope	ELISA	None, sequencing	HWYISKPQ		Chargelegue et al., 1998 [18]
Pentapeptide	5	1024	β-endorphin	ELISA	Recursive	YGGLL		Erb et al., 1994 [19]
Hexapeptde (EXXXPX)	16	512	gp120 Antibody	ELISA	iSURF	ESTRPM	None reported	Kerr et al., 1993 [20]
Pentapeptides	19	2 476 099	anti b-endorphin Ab	ELISA	None, sequencing	YGGFL	None reported	Lam et al., 1993 [21]
Hexapeptide	10	1 000 000	PK99H	ELISA	Positional scanning	EQFIPK	None reported	Wong et al., 1994 [22]
Pentapeptide	20	3 200 000	IL6 receptor	ELISA	Positional scanning	EFLIW	AAA	Wallace et al., 1994 [23]
Tetrapeptide	24	331 776	*S*-farnesyltransferase	E	iSURF	HWTD	NMR, MS	Boutin and Fauchère, unpublished
Tetramers	19	65 431	Trypsin	E	iSURF	xanthenyl(K/V/P/I)	ESI-MS	Carell et al., 1995 [24]
Tetrapeptide	24	331 776	MMP	E	iSURF	H-βal-Ahx-H[h]	NMR, MS	Ferry et al., 1997 [25]
Cyclic undecapeptide (SCXXSXPPQCY)	20	8000	Chymotrypsin	E	None, direct analysis instead	SCTYSIPPQCY	None reported	McBride et al., 1996 [26]
Cinnamyl tripeptide	5	125	Protein tyrosine phosphatase	E	None, radio frequency tagging	Cinnamyl-GEL	None reported	Moran et al., 1995 [27]
Tetrapeptide	22	234 256	HIV protease	E	iSURF	F-I-Sta-val[h]	None reported	Owens et al., 1991 [28]
Hexapeptide	20	64 000 000	Prohormone convertase 1 and 2	E	Positional scanning	Ac-LLRVKR-NH2	None reported	Apletalina et al., 1998 [161]
Pentadecapeptide (MAXXXXSXXXXAKKK)	15	2 562 890 625	Serine protein kinases	Es	None, sequencing	MAHHHRSPRKRAKKK	AAA	Songyang et al., 1994 [29]
Pentadecapeptide (MAXXXXYXXXXAKKK)	15	2 562 890 625	Tyrosine protein kinase	Es	None, sequencing	MAEEEIYGEFEAKKK	AAA	Songyang et al., 1995 [30]
Octapeptide	20	25 600 000 000	Protein kinase A	Es	iSURF	Ac-RAERRASI-NH2	None reported	Tegge et al., 1995 [31]
Heptapeptide (LXRASLG)	19	19	Serine protein kinases	Es	None, sequencing	LRRASLG	AAA, MS, Sequencing	Till et al., 1994 [32]
Tridecapeptide (RRLIEDAXYAARG)	19	19	Tyrosine protein kinase	Es	None, sequencing	RRLIEDAIYAARG	AAA, MS, Sequencing	Till et al., 1994 [32]
Tetrapeptide	28	614 656	*S*-farnesyltransferase	E	Positional scanning	wm-fcl-Gla[h]	AAA	Wallace et al., 1996 [33]
Heptapeptide	19	893 871 739	Protein kinase A	E	None, sequencing	SQRRFST	None reported	Wu et al., 1994 [34]
Pentapeptide	19	2 476 099	Protein kinase A	E	None, sequencing	RRYSV	None reported	Wu et al., 1994 [34]

Table 1. Continued

Tridecapeptide (XXXXXXHXXXXXX)	20	4×10^{15}	Phosphotransferase	E	iSURF	KKWHLRKR	None reported	Mukhija et al., 1998 [35]
Hexapeptide	20	64 000 000	HIV integrase	E	iSURF	HCKFWW	None reported	Puras-Lutzke et al., 1995 [36]
Tetrapeptide (CXXX)	19	6498	*S*-farnesyltransferase	E s	iSURF	CKQQ	NMR, MS	Boutin et al., 1998 [37]
Fmoc-Undecapeptide (FmocXAXXY*LIPQQG)[a]	8	512	Protein tyrosine phosphatase	E s	None, sequencing	EAEQY*LIPQQG[h]	AAA?	Cheung et al., 1997 [38]
Heptapeptide	19	893 871 739	p60 c-src	E s	None, sequencing	YIYGSFK	None reported	Lam et al., 1995 [39]
Heptapeptide (XYIXXXX)	19	2 476 099	p60 c-src	E s	None, sequencing	GIYWHHY		Lou et al., 1996 [40]
Decapeptide libraries	19	?	TGF-β receptor tyrosine kinases	E s	Positional scanning	KKKKKK(S/T)XXX	None reported	Luo et al., 1995 [41]
Tripeptide (E(C)X)	16	16	Glutathione *N*-transferase	E s	None, direct MS analysis	E(C)G (glutathione)	LC/MS	Wigger et al., 1997 [42]
Nonapeptide (XXY(NO2)XXXXXXK)	20	1 280 000 000	MMP9	E s	None, sequencing	Yes	None reported	Renil et al., 1998 [43]
Tripeptides	25	15 625	Vasopressin receptor	B	Orthogonal	Tic-Y-Nip[h]	None reported	Déprez et al., 1995 [44]
Hexapeptide	18	34 012 224	Opioid receptor	B	Positional scanning	YGGFMY	AAA	Dooley and Houghten, 1993 [45]
Hexapeptide	20 D-	52 128 400	Opioid receptor	B	iSURF	Ac-rfwink-NH2	None reported	Dooley et al., 1994 [46]
Derivatized tripeptide	17	289	Carbonic anhydrase II	B	None, MS analysis instead	CBS-LL-βal[h]	AAA	Gao et al., 1996 [47]
Hexapeptide	19	361	Endothelin receptor	B	iSURF	Ac-bhg-LDVIW-OH[h]	None reported	Neustadt et al., 1995 [48]
Dodecapeptide (GDGY*XXXSPLLL)	18	5832	SH2 binding domains	B	None, sequencing	GDGY*EEISPLLL[h]	AAA	Songyang et al., 1993 [49]
Dodecapeptide (GDGY*XXXSPLLL)	18	5832	SH2 binding domains	B	None, sequencing	GDGY*MEPSPLLL[h]	AAA	Songyang et al., 1994 [50]
Dodecapeptide (GAXXXY*XXXKKK)	18	ca. 34 000 000	SH2 binding domains	B	None, sequencing	GAEDDY*VEMKKK	AAA	Songyang et al., 1995 [51]
Hexapeptide (XXOOXX)	17	83 521	a.o. dsDNA binding	B	None	XXKYXX	AAA, Sequencing	Kramer et al., 1998 [52]
Pentapeptide	7	16 807	Streptavidine	B	None, direct MS analysis	Ac-HPQF	None reported	Youngquist et al., 1995 [53]
Pentadecapeptide (KEXYXZLAAXADAEK)	9	6561	HLA	B	None	Yes	Sequencing	Tana et al., 1998 [54]
Octapeptide	19	16 983 563 041	MHC	Cell-based assay	Positonal scanning	?	AAA, PSA, EMS	Udaka et al., 1995 [55]
Hexapeptide	18	34 012 224	Bombesin antagonists	Cell-based assay	Sequencing after controlled removal	AVGHLM	None reported	Jayawickreme et al., 1994a [56]
Nonapeptides	20	31 360	α-MSH antagonists	Cell-based assay	Sequencing after controlled removal	MPfRwFKPVN-H2	None reported	Jayawickreme et al., 1994b [57]
Tripeptide	96	221 184	MSH receptor	Cell-based assay	iSURF	wRL	None reported	Quillan et al., 1995 [58]
Heptapeptide (YGRGXXX)	19	1900	Platelet adhesion	Cell-based assay	None	YGRGDXX		Hortin et al., 1992 [59]
Tetrapeptide	24	331 776	IL5 receptor	Cell-based assay	iSURF	MDHL	NMR, MS	Boutin and Fauchère, unpublished

Table 1. Continued

Library type[a]	Alphabet[b]	Complexity[c]	Target[d]	Assay type[e]	Deconvolution[f]	Hit[g,h]	Analysis[i]	References
2/NON-PEPTIDES								
Pentamers (peptomers)	18	80 000?	Insulin receptor	ELISA	None, direct MS analysis	Yes (10 n*M*)	None reported	Ostergaard and Holm, 1997 [60]
Cyclic templates	29	24 389	Chymotrypsin	E	Positional scanning	Yes	None reported	Eichler et al., 1994 [61]
Trimers	9	810	Phospholipase A2	E	iSURF and positional scanning	Yes (5 μ*M*)	None reported	Wilson-Lingardo et al., 1996 [62]
Tetramers	20	400	Neurolysin	E	iSURF	Yes (4 n*M*)	None reported	Jiracek et al., 1996 [63]
Dehydrobenzopyran	62	7870	Carbonic anhydrase	E	Gas chromatography of tags	Yes	None reported	Burbaum et al., 1995 [64]
Acyl piperidine								
Monomers	?	20 000	Neurokinin NK1 and 2 receptors	B	None, encoding technique	Yes	None reported	Appell et al., 1998 [65]
Pentamers	20	146	Bradykinine B2 receptor	B	iSURF	Yes (80 n*M*)	MS	Chakravarty et al., 1996 [66]
Monomers	18	18	Carbonic hydrase	B	None, direct MS analysis	Yes	NMR, MS	Cheng et al., 1995 [67]
Piperazines	40	ca. 2500	Bradykinin receptor	B	None, direct MS analysis	Yes	None reported	Goodfellow et al., 1996 [68]
Dimers	80	1600	Neurokinin NK3 receptor	B	iSURF	Yes		Smith et al., 1994 [69]
Dimers	27	27	Antioxidant		iSURF	Yes	None reported	Kurth et al., 1994 [70]
Dimers/Trimers (peptoids)	24	ca. 4400	7TM receptor	B	iSURF	Yes (5 n*M*)	None reported	Zuckerman et al., 1994 [71]
Trimers	108	46 656	Antimicrobial	Cell-based assay	None, encoding technique	Yes	None reported	Silen et al., 1998 [72]
Trimers		4275	Antibacterial	Cell-based assay		Yes	None reported	An et al., 1998 [73]
Triazines	70	12 000			Not tested		MS, LC	Stankova and Lebl, 1996 [74]
Dimers	97	2001			Not tested		LC	Dankwardt et al., 1995 [75]

[a] Library type is the number of monomers with example of structures for biased libraries.

[b] Alphabet is the number of individual monomers present in the libraries.

[c] Complexity is the theoretical number of different structures found in libraries.

[d] Target is the molecular species towards which the library is aimed.

[e] Assay type covers: Es: search for enzyme substrate, E: search for enzyme inhibitor, B: radioreceptor binding assay.

[f] Deconvolution is the mean by which the best candidate structure has been found. iSURF: iterative synthetic unrandomization of randomized fragments.

[g] Hit is one example (often among several) of active compound found during the experiment.

[h] One letter code was used whenever possible. For non-natural amino acids, three-letter-code was used, as follows: βal: β-alanine; Ahx: aminohexanoic acid; Sta: statine; fcl: *p*-chlorophenylalanine; Gla: γ-carboxyglutamic acid; Y*: phosphotyrosine; Tic: Tetrahydroisoquinoline carboxylic acid; Nip: *p*-nitrophenyl-alanine.

[i] Analysis shows what kind of analytical characterization were attempted on the libraries.

Analysis), a large number of successfully used libraries reported in the literature, remain mostly uncharacterized. Obviously, the integrity of complex mixtures of thousands of components can only be partially demonstrated by analytical methods. Nevertheless, existing methods perform well enough to give a reasonable estimation of the success of the synthesis and the completeness of the sublibraries. Among them, as recently reviewed by Loo [76], mass spectrometry is certainly the method of choice in terms of both sensitivity and specificity for such characterizations. The recent applications of mass spectrometry in the field of combinatorial chemistry include the identification of compounds on solid-phase supports after detection of a biological activity or binding affinity of resin-bound ligands [77]. In the following sections, the discussion will be limited to the analysis of combinatorial mixtures of compounds (libraries) by these methods.

2.1. Low resolution mass spectrometry

The analysis of combinatorial libraries by mass spectrometry is generally performed by comparing the experimental mass spectrum to the theoretical molecular mass distribution. The choice of the mode of ionization is very important and some conditions must be fulfilled for reliable results to be obtained. Ideally, to avoid fragmentations and uninterpretable data, the ionization method should give a single and predictable ion, generally a once protonated or once deprotonated molecule, with a constant yield, although this criterion is never totally achieved experimentally. Amongst the possible ionization methods, electrospray (ESI) seems to be the most widely used [47,78–90], but liquid secondary ion mass spectrometry (LSIMS) [7,91] as well matrix-assisted laser desorption ionization (MALDI) can also give good results [53]. The major limitation of this simplest approach (ESI) is ionization efficiency which is strongly dependent of the chemical nature, hydrophobicity and basicity of the candidate molecule [80,87]. For example, with LSIMS ionization, the extreme basicity of arginine is responsible of the ten-fold overestimation of Arg-containing peptides in tetrapeptide libraries [7]. With ESI, where suppression effects are generally considered more effective than in LSIMS, an overestimation of the same order of magnitude was observed for basic diamino acid xanthene derivatives [81]. In the latter study, the equimolarity was better estimated in the negative mode and a combination of both the positive and negative modes allowed the diversity of libraries containing dozens of components to be demonstrated. The effect of the lack of one amino acid on some parameters of the mass distribution of peptide libraries was also investigated [92]. The mass shift between the average mass of a complete tetrapeptide library constructed by using a set of 20 amino acids (20^4 peptides, 381 different integer masses) and an incomplete one using a set of 19 amino acids (19^4 peptides, 381 different integer masses) was proportional to the difference between the mass of the excluded amino acid and the average mass of the whole set. Depending on the excluded amino acid the theoretical mass shift varied from -14 to $+15$ Da. On the other hand, if an average-weight amino acid, like Asp, is excluded, the mass shift is only 1 Da and cannot be reasonably estimated experimentally. The effect of an incomplete synthesis resulting in a mixture of tri- and tetrapeptides was examined in the same study. The authors concluded that the resulting shift of the mass average was appreciable only for a significant fraction of tripeptides. To illustrate this fact, we have calculated a simulated mass spectrum of a O1X2X3X4 library (24^3 tetrapeptides) for which the coupling step of X2 occurs with a low yield of 20%, resulting in a mixture of tripeptides O1X3X4 and a mixture of tetrapeptides O1X2X3X4. The poor quality of the mixed library is clearly reflected in the mass spectrum (Fig. 2). Even if low resolution mass spectrometry is limited for probing equimolarity or integrity of large libraries (e.g. cysteine oxidation), the mass spectra obtained with soft ionization techniques can be considered a good representation of the diversity of the library and pitfalls in the synthesis resulting in mass shifts of a dozen of Daltons (e.g. undeprotected peptides or truncated peptides [88]), will easily be detected even for large peptides libraries.

2.2. High resolution mass spectrometry

Fourier transform ion cyclotron resonance mass spectrometry (FT–ICRMS) is the most powerful mass spectrometry technique in terms of resolution

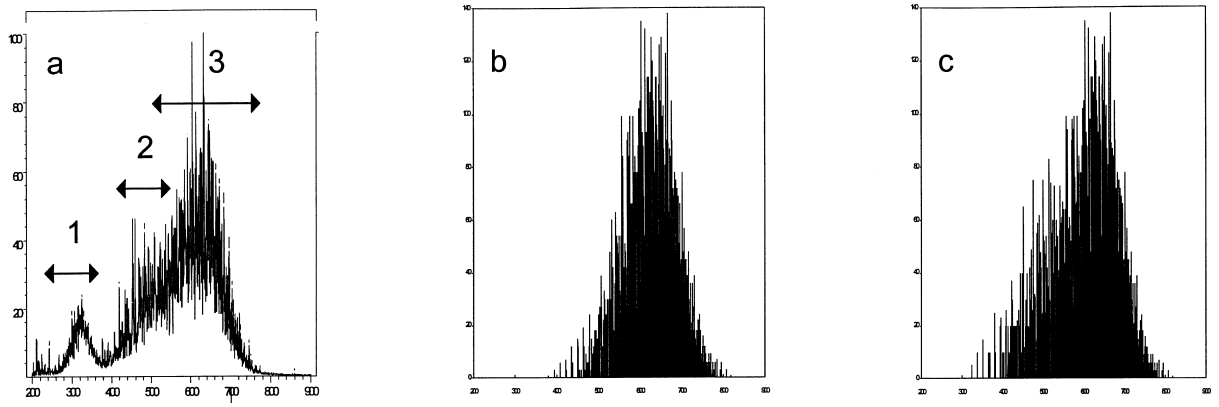

Fig. 2. Combinatorial library mass spectrum. a: Experimental ESI mass spectrum of the failed synthesis of library O1X2X3X4-NH_2 (O1=*p*-nitrophenylalanine, and X any of a set of 24 common amino acids) showing three areas; 3 – mass range of protonated tetrapeptides O1X2X3X4-NH_2; 2 – mass range of protonated tripeptides O1X3X4-NH_2; 1 – mass range of doubly protonated tri- and tetrapeptides. b: Calculated m/z distribution of protonated peptides for an equimolar and complete tetrapeptide library. c: Calculated m/z distribution of protonated peptides for a mixture of tri- and tetrapeptides libraries assuming that the coupling step of X2 occurs with a yield of 20%.

and mass accuracy. The specificity of mass spectrometry can be greatly enhanced, by increasing resolution which allows to check in detail mass diversity and degeneracy of libraries. Blom [93] has calculated that for a 12×96 non-peptide library a mass accuracy of 5.2 ppm is necessary to reach the maximum specificity (i.e. the fraction of components which can be separated by mass spectrometry). For a more diverse 12×96×6 library, for which the maximum specificity is around 33%, a mass accuracy of 4.2 ppm is needed. In both cases, the maximum reachable specificities decrease rapidly with decreasing resolution. The benefit of a resolution giving a mass precision of 100 ppm is not evident compared to a unit mass resolution. For peptide libraries, for which degeneracy is higher (due to the redundancy of masses and elemental compositions of amino acids), the highest reachable specificities are lower than a few percent. Winger and Campana [90] have shown that the relatively high resolution obtained only on magnetic sector or FTMS instruments was not only sufficient to separate lysine from glutamine components (which differ by 0.036 mass units at m/z 945) in a small 19 components SIIN-X-EKL library, but also to unambiguously identify the glutamic acid component at m/z 946, which would be indistinguishable from its ^{13}C isotopic peaks. Very high resolution (which can only be reached with a FTMS instrument) is necessary to resolve some multiplets from larger libraries containing 19^3 tripeptides [88].

2.3. Tandem mass spectrometry (MS–MS)

Another way to increase the specificity of mass spectrometry is the use of tandem mass spectrometry. When predictable and general reactions of ions occur, they can be monitored by product or precursor ion, or by neutral loss scanning, to specifically detect a particular class of reacting ionized compounds. For example, protonated arginine-containing peptides undergo an elimination of the C-terminal residue. Consequently, a scan of neutral C-terminal residue losses will detect all peptides containing both arginine *and* the given C-terminal residue. This method has been successfully used to control the quality of O1O2X3X4 tetrapeptides libraries (24^2 components) [7]. Other MS–MS experiments have been used for library analysis [80–83]. Residual protected peptides can be specifically detected in mixtures of 48 O1O2O3O4O5X6X7X8 octapeptides [87]: neutral loss scan of 56 Da reveals the presence of Boc protected peptides, whereas neutral loss scan of 309 Da shows Pmc-containing peptides. Finally, all tritylated peptides will be detected by a precursor ion scan of m/z 243 (triphenylmethylium cation). In some cases, missing

compounds which were not detected by a direct ESI mass spectrum, could appear with MS–MS experiments [83]. However, as noted by Blom [93] for non-peptide libraries, the prospect of reliably predicting MS–MS product ions in libraries made up of thousands of diverse components (one scaffold with three diversity points) seems slight. MS–MS is more realistically to be considered a tool for extracting structure-related families of compounds in libraries of moderate size.

2.4. Liquid chromatography (LC) and capillary electrophoresis

HPLC is certainly of high interest for the characterization of libraries although it is essentially limited to small libraries (e.g., 19 nonapeptides G*X*ATPQDLNT where *X* is one of the 19 natural amino acids, cysteine omitted) [94,95]. Of course, the diversity of libraries containing hundreds to thousands of compounds is only roughly estimated by HPLC alone [96]. With CE, chain deprotection and incorporation of amino acids can be checked for relatively large peptide libraries. Hence the 24^3 peptides of the library O1X2X3X4 (net charge from -5 to $+5$) could be classified into different classes depending on the theoretical net charge at basic or acidic pH values and a good correlation was found between peak integration and theoretical population of each class of peptides [7].

2.5. LC–MS and LC–MS–MS

The two- or three-dimensional time/mass spectral data provided by coupling chromatography and mass spectrometry is obviously the more promising way to characterize libraries. For small libraries containing less than 100 components, the sensitivity and specificity of the mass spectrometric detector combined with the separation efficiency of reversed-phase HPLC allow the retrieval of all expected ions as well as the detection of non-UV absorbing impurities [82,86]. By plotting reconstructed ion chromatograms, the number of compounds, assuming they are separated, of any given mass can be counted even by coelution of compounds of different mass. This approach is not totally free of artefact peaks: peaks due to isotopic contribution of lower m/z and peaks due to ionic species other than $[M+H]^+$ (adduct ions, multiply charged ions, etc..) are expected. For sensitivity reasons, scanning over a complete mass range seems to limit LC–MS to the analysis of moderate size libraries of typically less than 100 components. The selected ion monitoring (SIM) technique, where just one or few masses are recorded, can increase sensitivity by two or three orders of magnitude. In a O1X2X3X4 ($24^3=13\,824$ tetrapeptides) library [85] a peptide known to be unique at a given mass (i.e. the heaviest peptide) and representing theoretically 4.6 pmol, could easily be detected with a satisfactory signal-to-noise ratio in agreement with the ratio obtained with an external standard. Consequently, this LC–ESI-SIM approach can be efficiently used to evaluate the diversity of medium to large libraries containing up to thousands components and the limitations are mainly set by the chromatographic performances.

2.6. Practical considerations

From a pragmatic point of view, analyses of libraries can be conveniently divided into a first category devoted to the assessment of chemical synthesis and a second one aiming at the description of the library composition.

FTIR spectroscopy allows the progress of some organic reactions to be followed on a single bead [97]. Similarly, magic angle spinning nuclear magnetic resonance (MAS-NMR) spectroscopy is very powerful to give in a non-destructive way, good quality HMQC and TOCSY data for compounds attached to polymer supports [98]. Direct, real-time monitoring of organic reactions can also be conveniently done by MALDI mass spectrometry especially when photocleavable linkers are used [99]. These analytical performances are also potentially useful for the identification of ligands on beads found to bear an active species in a biological test.

Both the molecular diversity and the equimolarity of a given library are parameters of interest which can be explored by analytical methods. Mass distribution is easily described by conventional soft ionization mass spectrometry (see Metzger et al. [87], Boutin et al. [7] and Fig. 2) and the presence of common structural and functional groups can be investigated by tandem mass spectrometry. For large

libraries, selective ion monitoring (LC–ESI–SIM, vide supra) offers a good quality control of the diversity of the library. Equimolarity of the components of a given library is generally more difficult to assess. LC–UV can rarely be used due to large variations of the absorbances of the individual components. 2D–NMR was used [7] to estimate the relative amounts of each amino acid in the mixture, thus also giving a record of the mean incorporation of each residue in the peptide sequences. However, this very informative procedure, which could be in principle used for other non-peptide homogenous collections, is too complex to be applied on a routine basis for each new library.

On the whole, it can be seen that a pragmatic strategy has to be adopted in each particular case and that a combination of existing analytical methods, among which mass spectrometry plays a major role, most often leads to an extensive characterization of the synthetic libraries.

3. Biological analysis

The success of the discovery of new pharmacological leads depends on several factors, mainly on the quality of the libraries (which relies on synthesis and analysis), the quality of the assay (which relies on specificity of the target) and obviously, on the molecular diversity of the screened libraries. While the synthetic tools for the generation of combinatorial libraries and the merits and limitations of the various deconvolution techniques have been reviewed (see, in particular, Krchnak and Lebl [100], Felder and Poppinger [101] and Fauchère et al. [11]), only a few authors have delt with the comparative effectiveness of the biological assays in library screening. It is obvious, though, that problems encountered while screening several thousand compound mixtures (Fig. 1) in biological assays, are rather different from those encountered in conventional radioreceptor binding or enzyme assays. A number of representative examples are gathered in Table 1 which may serve as a basis for the evaluation of the most successful strategies in lead identification. In the following sections, these studies shall be briefly analyzed in terms of the used tactics of deconvolution, target and assay types and achieved molecular diversity.

3.1. Deconvolution

3.1.1. Iterative SURF

Finding the needle in the haystack is made possible by the effectiveness of the various deconvolution techniques among which the simplest and most robust one is the iterative procedure first proposed by Furka et al. [5] and popularized by Houghten et al. [6]. This procedure, labelled SURF (synthetic unrandomization of randomized fragments), has been concisely described by Freier et al. [102] as follows: 'SURF deconvolution begins with synthesis of a nonoverlapping set of mixtures by incorporating a unique monomer at a common position of each subset. The subsets are tested separately and the one with greatest activity is identified. A second set of compound mixtures is prepared with each subset containing the fixed monomer showing greatest activity from the previous round. In addition, another position is fixed with each of the unique monomers to give another set of subsets. The complexity of the mixture is reduced and the process is repeated until a unique molecule is identified.' As can be seen from Table 1 (column 'Deconvolution'), a large percentage of combinatorial work has been done using this strategy in which the whole system becomes simpler while the deconvolution progresses, and as such, can be described as a purification procedure. Already after a few years of use, this deconvolution system has been extended to non-peptide libraries [63,66,69–71].

From the point of view of biological evaluation, a striking feature is the high complexity of the mixtures to be tested in the first round, which could question the reliability of the results. However, during the search for enzyme inhibitors using tetrapeptide libraries, we found clear-cut answers in the first round of screening, as can be also seen from various published results [33,46]. The most likely working hypothesis for such clear results in highly complex mixtures, is that 'families' of analogues of active compounds are acting in a cooperative way to inhibit the activity. This cooperativity led to apparent massive inhibition, while deconvolution of these active sublibraries often led to compounds with only

marginal potency. This apparent discrepancy between theoretical expectations in terms of inhibitor potencies calculated from results obtained during the first round of screening, is indeed explained by this apparent cooperativity. The frustration from apparently 'poor' results should be compensated by the fact that the structures discovered are unlikely to be found *ex nihilo* and constitute a solid and original basis for new classes of pharmacological agents (see Ferry et al., for example [25,103]). In our hands, the deconvolution systematically led to novel inhibitory structures, despite potencies that were often inferior to those extrapolated from the first round of test (a feature also described by Dooley et al., [46]). The possible occurrence of false positives did not abolish the success of deconvolution!

This particular deconvolution system can be used for the search of enzyme substrates, as for example, of the *S*-farnesyltransferase [37]. For substrate search, the capacity of measuring the presence of one or several substrates in the initial mixtures is difficult due to technical problems. Indeed, in transferase assays, the substrate is enzymatically modified as opposed to inhibitors which only modulate the enzyme activity. The search for substrates is made possible by indirect methods of measurement, for instance, of the disappearance of the cosubstrate of a transferase reaction. In several sets of separate experiments with different transferases in which the effects observed in the first round of deconvolution (tetrapeptide libraries of less than 6000 compounds), had poor amplitude, statistical significance was ascertained by increasing the number of independent measurements. Based on at least 20% difference in consumption of cosubstrate compared to the control, identification of new substrates was achieved [37]. It therefore appears that the most important part of the screening process of libraries (mixtures) is the technology used for the very first round, even if the assay is repeated several times. However, working on about 20 targets for the last 3 years, we failed to obtain ligands on only two occasions, although in some successful cases, the obtained ligands were of modest potency [25].

Table 1 gathers some representative publications dealing with the biological screening of libraries. Fourteen, out of 67 examples, used the iterative deconvolution system described above (SURF) [5,6], a rather low score when compared to the later developed techniques such as the positional scanning approach or the on-bead assays followed by sequencing. Nevertheless, SURF remains the most trustable technique for large libraries which result from the sequential assembly of building blocks, such as polyamide or polynucleotide libraries.

3.1.2. Positional scanning

Positional scanning [104] is a non-iterative procedure where for each position of the oligomer sequence, a series of mixtures is synthesized with a different monomer in the fixed position. Each of the mixtures is tested separately and the lead molecule is deduced by selecting the monomer from the most active mixture from each position set. In principle, a single round of screening is required to define the most active molecule. In fact, both theoretical [102] and experimental [45,161] evidences have shown that peptides obtained according to this procedure are not necessarily the most active structures. However, in practice, since several building blocks (amino acids) can give comparable activity in a given position, the synthesis of a number of oligomers (peptides) may be necessary to obtain a highly active lead (e.g., if two amino acids give positive responses at each position in a hexapeptide library, 64 peptides will have to be synthesized and screened). In addition, since the contributions of each position to the overall activity are not independent, there is no reason to think that the most active oligomer will be found. This fundamental problem is only partially reduced when a two-positional scanning is performed [104] or when a domino overlap of the 'active' segments is followed [105]. Another concern about the unavoidable coupling of mixtures of incoming amino acids has been raised [79], and demonstrated to lead to uneven representation of the goal products, due to different coupling kinetics. Nevertheless, positional scanning is still often used for the benefit of starting all the syntheses simultaneously and of avoiding the mix and divide steps which are not available in most of the automated apparatus for parallel synthesis, or for the purpose of synthesizing very large libraries [15] which could not be obtained without coupling of mixtures. Positional scanning has led to interesting data on protein/protein interactions reported by Songyang et al. [49–51], a break-

through in the understanding of the phosphopeptide/SH2 domain interaction, as a key step in second messenger biochemistry. Nevertheless, because the completeness of the library obtained by mixture on mixture coupling in solid-phase synthesis is far from the theory [79], there is no guarantee that all the possible peptides interacting with a given structure will be documented. In other words, other sequences might be discovered, beside those described by Songyang et al. [49–51], which interact in a similar way.

3.1.3. On-bead screening of one-bead-one-peptide libraries

When, in contrast to soluble libraries, biological detection is performed with the ligand still attached on the solid support, deconvolution of one bead-one peptide libraries is avoided (for reviews, see Refs. [106,107]). In most cases, no experimental evidence demonstrates the completeness of the libraries. Furthermore, this technique requires a methodology to identify the lead by eye either under the microscope in a ELISA-type assay [10], by death cell in an antimicrobial assay [72] or after impression of a photographic film [40] by a labeled enzyme substrate. While some assays are relatively easy to use under those conditions, as for protein kinase substrates [34,39], they are less easily performed when inhibitors are searched for. Indeed, immobilized peptides can interact with the enzyme and inhibit the reaction, thus leading to an observed diminution of the net catalytic activity compared to a control reaction. However, in contrast to soluble libraries where a programmed deconvolution retrieves the active ligand, it is almost impossible to find a way to positively identify a bead carrying the enzyme-inhibiting peptide. The only cases for which that might be feasible are when the enzyme is finally immobilized on the bead carrying the active peptide [47,106]. Once again, a large number of successful studies using ligands immobilized on beads have been reported in the literature (reviews by Lebl et al. [106] and Lam et al., [107]). This approach has some clear advantages, among which the fact that components of the library are spatially separated, and therefore all the peptides can be tested at the same time, under the same conditions. Therefore, compounds with completely different motifs can be identified at once from a single assay. On the other hand, the main challenge linked to this technology is the fact that once recognised, the positive bead should be analyzed to allow the structure of the active compound to be determined, a problem solved by sequencing only for natural peptide (or oligonucleotide) sequences.

3.1.4. Spot synthesis and screening

Frank described the synthesis of libraries on a cellulose-based matrix [108]. This technique can be used for the synthesis of both collections of individual peptides or peptide libraries. For the latter purpose, one or several mixture on mixture coupling step(s) are involved, raising again the problem of variable coupling kinetics of the incoming residues. The spatial deconvolution used in this methodology efficiently retrieves individual sequences in large libraries. Indeed, despite the apparently strong degeneration of these libraries (several millions of individual compounds, see Ref. [108]), the only feature that is really looked at is a fixed region of usually *two* [109], sometimes *six* residues [110], all constructed from 17 amino acids building blocks. The library is arranged in such a way on the paper sheet, that columns represent one fixed position while rows represent the other [108] and therefore, positive spot identification in a relevant assay reveals the active sequence immediately. A weakness of this powerful technique is that it is impossible to analytically determine the nature of the mixture spotted on the paper sheet and that the success of the synthesis has to be taken for granted despite already critized mixture on mixture coupling step [79]. Apart from protein kinase A [31], this technique might be conveniently applied to the discovery of substrate consensus sequences of many protein-maturating enzymes such as metalloproteases, acyl-transferases or *N*-myristoyltransferases for which long sequence substrates (octa- to decapeptides) are required.

3.2. Assays and targets

The success of screening process of complex (peptide) libraries depends not only on the presence of potent ligands in the library, but also on the biological assay system. As a rule, the first screening test should be experimentally easy to perform and

the biological activity clearly detectable. Such conditions are generally met in ELISA (UV absorbance detection), SPA (fluorescence), radioreceptor assay (radioactivity). More importantly, the measured effect should be clearly assignable to the molecular target aimed at. From this point of view, assays performed on an isolated enzyme are highly specific and likely to reliably identify inhibitor or substrate ligands when used for the deconvolution of large libraries. The same enzymatic assays, when performed on whole cells [111] are more difficult to run even if the molecular target is the same, due to the complexity of the cellular system. Potential problems may arise from cell penetration, access to the target or interaction with other cellular components. The search for antagonists of G-protein-coupled receptors is generally performed in binding experiments in which subtype and species specificities can be insured by using cloned receptor preparations. The search for receptor agonists can sometimes occur on acellular membrane preparations in which the production of second messenger [112] or the activity of the GTPase [113] are estimated. For library screening, this identification of agonists would be more convenient using simple in vitro experiments than bioassays on isolated tissues (e.g. in which the stimulation of contractions of the tissue is measured) or on cultured cells, since several synergic mechanisms may mediate the observed effect, thus leading to ambiguous interpretation of the results, and to possible false positives during deconvolution. Finally, selection of active compounds by injection of complex libraries in the whole animal (rabbit) followed by the observation of a pharmacological effect (decrease of blood pressure), although suggested, seems unreasonable.

3.2.1. Catalytic targets: the enzymatic assays

When using the measurement of the activity catalyzed by pure or partially purified enzymes for the deconvolution of large libraries, there is a high probability to find a hit substrate or inhibitor [24,25,29,30,34,37], although some of these hits may show low activity, typically in the millimolar range.

3.2.1.1. Substrates. Thanks to the combinatorial methods, it is now possible to do systematic and comprehensive studies of the substrate specificities of enzymes. Such studies requiring the use of numerous individual substrate compounds (see for examples: UDP-glucuronosyltransferases [114], *N*-myristoyltransferase [115], tyrosine protein kinase [116], *S*-farnesyl- and geranylgeranyltransferases [117,118]), are often impaired by the high cost and the difficulty in obtaining or synthesizing discrete series of individual compounds. Even data on extensive series of tyrosine protein kinase [119] or *N*-myristoyltransferase [120] substrate sequences did not lead to the knowledge of consensus sequences [essentially because the structure of the substrates studied were non-continuous]. Combinatorial libraries can now afford complete collections of sequences and provide essential and continuous information on enzyme specificities [37]. The experimental limitations of specificity studies comprise the use of natural mixtures of isoenzymes (as opposed to purified single proteins). Indeed, the enzymology of isoenzymes acting together at the same time on a common substrate is too complex (e.g. UDP-glucuronosyltransferases, [121]), to insure a safe deconvolution process.

For screening enzyme substrates among a population of candidates (essentially peptide structures), the following approaches seem to be usable. (1) Immobilization of the families of substrates (as in Refs. [31,34,35]), either on beads or on cellulose sheets, but followed by resynthesis and deconvolution or in some cases, the bead bearing the enzymatically-transformed substrate can be identified by direct sequencing, for instance, as a result of the transfer by a protein kinase of a [^{32}Phosphate] from [^{32}P]-γ-ATP onto the immobilized substrate, the bead becomes radiolabeled, can be spatially localized (see Ref. [40]) and sequenced. Other examples could include the enzymatically-catalyzed transfer of [^{3}H]acetyl moiety from acetyl-CoA by acetyltransferases, [^{3}H]farnesyl or [^{3}H]geranylgeranyl residues from their pyrophosphate counterparts.... The easiest techniques include peptide-modifying transferases. Other, more sophisticated techniques might be used such as biased immobilized libraries comprising two fluorescent residues able to quench each other and borne by a family of peptide potentially substrate of a given protease (see Ref. [122], for example). Suppression of the quenching effect and appearance of the fluorescent signal, permit to iden-

tify the bead carrying the relevant substrate. Alternatively, libraries N-terminally derivatized by a fluorophore would also, after enzymatic cleavage, permit either the identification of the active bearing bead or of the active-containing sublibrary during the deconvolution process.

(2) Indirect measurement of the reaction, for instance by estimating the amount of cosubstrate (in a transfer reaction) comsumed during the catalytic reaction and the use of this assay to find new substrates (e.g. *S*-farnesyltransferase, [37]). Other examples might include enzymes using coenzyme A derivatives as cosubstrates of a transfer reaction. Generally, cosubstrate measurement can be used as a marker of an enzyme-catalyzed reaction, as long as the decrease in cosubstrate can be specifically linked to the catalytic reaction. Cosubstrate disappearance should be easily followed by any robust assay. Furthermore, specificity studies could be conducted using xenobiotic metabolizing enzymes such as UDP-glucuronosyltransferase, glutathione-*N*-transferase, cytochrome P450, sulfo-transferases . . . , by using non-peptidic structures immobilized on beads. Indeed, most of these enzymes are using cosubstrates, the consumption of which can be again easily followed.

As reported in Table 1, the peptide specificity studies were conducted mainly on transferases [37,42], kinases [29,30,39–41] and phosphatase [38].

3.2.1.2. Inhibitors. An important purpose of combinatorial libraries is to provide the pharmacologist with vast numbers of compounds to be screened on enzyme activity [123]. Using a mixture of compounds might be a source of problems because several inhibitors are potentially present at the same time. The kinetics of inhibition of an enzyme activity by several inhibitors as described by Chou and Talalay [124,125], shows a high degree of complexity and does not follow a classical Michaelis–Menten law. Therefore, IC50 or *K*i determinations using compound mixtures at early stages of deconvolution are elusive. Those determinations sometimes attempted in order to predict the potency of the inhibitor in the mixture of tested compounds assuming only one compound is responsible for the inhibitory activity of the mixture and lead to too high apparent *K*i values. Similarly, the differential screening of libraries on several targets in order to obtain selective inhibitors might be a difficult task. Selectivity should be gained using a sequential approach, first by selecting a 'main' target and then, once a compound is identified, by using secondary targets against which poor activities of the selected compound(s) will be preferred as reported for selected panels of kinase inhibitor studies [126,127].

Although technically amenable to all kind of combinatorial libraries (immobilized as well as soluble), inhibition studies will be preferentially but not exclusively conducted using cleaved soluble libraries, especially because immobilized (on beads or on sheets) libraries-mediated inhibitory activity could be impaired by the spatial hindrance of the support. For bead-released libraries, the screening techniques used are similar to those used with soluble libraries. Further possibilities, though, are offered with double-cleavable linkers as reviewed by Lebl et al. [106].

When using immobilized inhibitors, it is crucial that the binding of the enzyme on the inhibitor on the bead is strong enough to stabilize the enzyme/inhibitor complex (as described for carbonic hydrase, [68]). Only then might the bead bearing the active inhibitor and the enzyme be positively recognized by an antibody against the enzyme and the inhibitor peptide immobilized be sequenced. Obviously, (non-spatially arranged) one-bead-one-peptide libraries cannot easily be used for inhibitor studies, as also suggested by Lam's review [107] where one only such example is given [128]. Although not yet reported in the literature, immobilized libraries, spatially arranged as in the SURF methodology, could also be used for inhibitor discovery implying a deconvolution process.

Inhibition studies have been conducted mainly on proteases [24–26,28], tranferases [33,35], phosphatases and kinases [27,31,32,34].

3.2.2. The non-catalytic targets: the binding assays

Besides antigen–antibody binding (the most frequent and probably most sensitive technique of binding), and radioreceptor binding assay, the affinity between proteins such as SH2 domain/phosphorylated peptide interactions have also be consid-

ered for the deconvolution of peptide libraries (Songyang et al., [49–51]).

3.2.2.1. Receptor binding. Most receptors are dynamic transmembrane proteins with binding capacities and no catalytic (transforming) capacities. G-protein-coupled receptors are frequent targets of pharmacological screening. Surprisingly, there are only a few reported examples of such a screening with peptide combinatorial libraries [43,45,46,48,129]. Cloned receptors of interest, expressed at an artificially high concentration (>1 pmol/mg of protein) will be preferred over naturally expressed ones (e.g., using a 'natural' source of EGF receptor overexpressed in A431 cell line, or the cell line KANT-S as source for neuropeptide Y receptor subtype 2). Receptor binding assays will not distinguish agonists from antagonists. The situation is less marked for enzymes, where the rare activators should not impair the discovery of potent inhibitors. However, in displacement assay, there is no problem in using a mixture of compounds (virtually agonistic as well as antagonistic compounds). Indeed, either of them will be selected of the basis of a shift of the receptor ligand binding to the contrary of at least one claim in the literature [130]. As a proof, we studied the behaviour of artificial compound mixtures in a binding assay. In brief, we chose 50 inactive compounds (IC50>10^{-5} *M*) issued from our HTS program on melatonin receptors. We selected five active compounds with a potency on the same receptors ranging from 10^{-8} to 10^{-12} *M* (IC50). Those compounds were individually mixed with an equimolar pool of the 50 inactive compounds, leading to five pools of 51 compounds each. These pools were tested using a displacement binding assay with [^{125}I] 2-melatonin from the cloned human mt1 receptor. The results clearly led to IC50's of the same order of magnitude than the pure compound, suggesting that within the limits of these experiments, active compounds can be identified by deconvolution among a mixture of inactive chemicals (J.A. Boutin, C. Lahaye, J.P. Nicolas, unpublished observations). Likewise, this is due to the nature of the binding measurements which are highly selective and the radiochemicals used, potent ligands with *K*i in the low nanomolar range. Library-based techniques combined to radioreceptor assays therefore provide a powerful tool to discover new receptor-binding entities.

3.2.2.2. Protein–protein Interactions. More and more evidences demonstrate the key role played by protein–protein interactions in biochemical processes. Examples include phosphorylated protein interacting with SH2 domain [131], the polyproline/WW module [132], seven transmembrane domain receptors and G-proteins [133,134] receptor homodimerization [135,136] or heterodimerization [137], as well as a cascade of events forming the third messenger pathway downstream receptors as described in many reviews (see for example Ref. [138]).

The use of peptide [139,140] or of protein (arrestin, [141]) ligands for identifying these pathways has started to be described in the literature. Songyang et al. [49–51] published a series of impressive studies dealing with phosphopeptides recognized by SH2 domain-bearing proteins in which they showed an interesting and general way to use fusion proteins to screen directly the phosphopeptide libraries.

The target protein is cloned and expressed as a protein fused at the C-terminus of glutathione *N*-transferase. The protein–protein association is measured by specific antibodies against the second protein. Peptides impairing this protein–protein interaction lead to a decreased antibody recognition of the second protein. Another conventional SURF deconvolution step is required.

The immobilized protein can serve as an affinity chromatography matrix for the peptides applied on to the column. The retained peptides are identified by sequencing . It does not seem possible to easily use peptide libraries immobilized on beads in these type of experiments.

3.2.2.3. Antigen–antibody interactions. The antibody recognition of an antigen is the most powerful molecular detection method to date essentially due to the high specificities and affinities observed. The pioneering work on peptide libraries of Geysen et al. [9] was indeed dealing with this technique. Since libraries are mixtures of low concentration compounds, to fish out a single, specifically recognized

structure requires a powerful technique. Many examples of such studies have been reported in the literature ([6,15–22,110], see also further references in [107]), with the aim of detecting non-contiguous amino acid sequences that form the antigen recognized by monoclonal antibodies.

3.2.3. Cell-based assays

Cellular-based assays are easy to use with mixtures, as long as cell death is the parameter measured in such experiments, as in cytotoxicity [72,142]. Other cell-based assays in which hormone or cytokine-mediated cell survival or induced endobiotics production are estimated, are difficult to run, poorly reproducible, probably due to active compounds acting by complex unknown mechanisms. A number of individual enzymes or receptors may be involved in a single functional chain of events in the whole cell. For instance, when the second messenger production or other effects downstream the receptor activation are the measured parameters, the library components might interact with the receptor itself (at various sites, with various effects), its coupling with the G-proteins, the signal transduction of G-proteins to their respective partners (such as *ras*, GAP, SFT, *src*, MEK, MAPK, MAPKK, etc.) and with the ultimate protein signal as for example the nuclear protein *c-fos*.

Nevertheless, functional assays of α-melanotropine (MSH) or bombesin receptors such as the pigment granule aggregation assays run on melanophore cells in culture have been described [56–58]. Change in color being the measurement parameter, these experiments led after the deconvolution to a tripeptide, trp-Arg-Leu-NH_2, [58], or Met-Pro-phe-Arg-trp-Phe-Lys-Pro-Val-NH_2 [57] with remarkable antagonist activities. Examples are given when topic treatment of the animal (xenopus) skin with these peptides can be read in vivo [58].

After transiently expressing the bombesin receptor into melanophore cells, the functional signal mediated by the bombesin receptor was then linked to the melanosome translocation [143] and this model used for the deconvolution of heptapeptidic agonists such as Ala-Trp-Val-Gly-His-Leu-Met-NH_2 [56]. Two other reports deal with peptide library-mediated analysis of the structural requirements for peptide binding to the major histocompatibility complex Class I. The assay involved cell treatment with the library for 24 h., followed by the treatment of the cells with an FITC-labeled antibody (B8.24.3) recognizing the conformationally intact K^b protein. Such an analysis was done by flow cytometry and led to the discovery of Arg-Gly-Tyr-Val-Tyr-Gln-Gly-Leu peptide [55]. Further studies using a similar cell-based assay were reported [144].

In spite of their complexity, it is obvious that such assay systems are practicable and have already led to new active compounds. The key element of those cell-based assays seems to be a rigourous assessment of the system.

4. Diversity covered

As already stated in the present review, combinatorial chemistry has revolutionized medicinal chemistry by speeding up the generation of molecular entities among which a biological test will select leads as potential drug candidates.

4.1. Libraries for lead generation

Libraries for lead generation have been described from dipeptides [145] to pentapeptides [54] containing from a few dozen to trillions [15] of compounds. Several reasons are in favour of a limitation of the sequence length to the hexapeptide size in peptide libraries. Firstly, this length is believed to be sufficient for proper epitope recognition [146], at least for a free peptide in solution (as opposed to a bead-linked peptide). Secondly, shorter segments are recognized by G-protein-coupled receptors, as suggested by Ariens [147], or are convenient leads (substituted dipeptides) for chemical optimisation [145]. Thirdly, calculations based on a Poisson distribution of the amount of resin needed to insure proper representation of each individual peptide after the mix and divide procedure, result in gram quantities in each reactor for hexapeptides, while unpracticably higher quantities of resin would be needed for hepta- or longer peptides [13]. Those libraries are mostly constructed with proteogenic amino acids (see for example Ref. [15,36,23,43]) or with additional non-natural amino acids [33,37]. Since obviously, some of the structural information encoded by

the natural amino acids is redundant, shorter selections based on analysis of their principal components can be used, sometimes even including non-natural amino acids [22,148,149]. Using this bias, one can build up larger libraries containing less letters but longer words (e.g. eight amino acids would allow octapeptide libraries to be reliably obtained on gram amounts of resins per reactor (see Refs. [13,106] for calculation).

4.2. Biased libraries for lead optimisation

Biased libraries for lead optimisation generally bear a key pharmacophore residue around which the rest of the library is constructed. Examples of constant building blocks are tyrosine [39] or serine/threonine [31,34,41] for protein kinase targets, phosphotyrosine for protein/protein interaction [49–51] and for tyrosine phosphatase studies [38], or the cysteine in CAAX box-bearing substrates of *S*-farnesyltransferase [37]. Similarly, other libraries have been built around a cinnamyl moiety [27] or the phosphonic acid derivatives [63]. Libraries of cyclic peptides [61,150] are also convenient means to look for conformationally restricted ligands which could serve as models for the design of peptide mimetics. Finally, a number of libraries have been constructed on a common centroid scaffold, such as triazine [74], purine [126], or piperazine [68] for which the corresponding expertise in organic chemistry is available.

Despite the great progress achieved in the production of large numbers of compounds by combinatorial or parallel synthesis, the question of the coverage of the available chemical space is increasingly being discussed. It has been demonstrated that the exploration of 3D-chemical space is easily obtained with peptides of at least the hexapeptide size, despite the rigidity of the peptide bond [151]. A comparable conformational diversity is difficult to achieve with heterocyclic libraries such as benzodiazepine, hydantoin or steroid libraries. In addition, the chemistry of the amide bond formation being rather simple, the synthesis of peptides on solid-phase very efficient and the diversity of the functional groups rather high, peptide libraries are still an attractive source of ligands for new biological targets. When leaving the peptide field for obvious reasons of proteolytic degradation and poor bioavailability, the evaluation of the molecular diversity of compound collections has become current practice, in order to reduce the number of samples and of biological assays and still keep an optimized coverage of the functional and conformational molecular diversity. A number of methods have been described to characterize the molecules in real or virtual libraries on the basis of 2D-fingerprints, 3D-pharmacophores or other classical lipophilic, steric and electronic properties [152–155]. Selection of representative molecules on this basis, is likely to enlarge the diversity covered by the same number of molecules, compared to a random selection.

5. Conclusions

As stated by many authors, the immense development of combinatorial chemistry opens new research areas, and saves time and money in the finding of new leads. It also stimulates the development of upcoming analytical techniques such as LC–NMR [156–158], solid-phase HR-MAS NMR [98]. . . . Combinatorial chemistry has certainly opened new avenues for new concepts and new approaches of analysis as discussed by Czarnik [159]. Whether new drug types or new therapeutic areas will also result form this technology remains to be demonstrated (for larger discussion, see Myers [160]). We believe that combinatorial chemistry will nicely complement the currently available tools in biological research, including the techniques involving large numbers of unique compounds (see Table 2). Besides extending considerably the synthetic means to generate molecular diversity, combinatorial chemistry has brought new challenges for the scientist. The analytical chemist has been compelled to face the qualitative and quantitative evaluation of complex mixtures or the structural elucidation of 100 pmol amounts of ligands on a single bead. The biochemist is now confronted with large numbers of samples produced by parallel synthesis which requires high throughput screening strategies, or with the testing of mixtures generated by combinatorial synthesis which call for robust deconvolution procedures. While library screening is of little value to establish structure-activity relationships, among the library congeners,

Table 2
Comparative qualities and defaults of libraries obtained from combinatorial or parallel syntheses

	Combinatorial synthesis	Parallel Synthesis
Synthesis	In reactors (typically 24 to 36) (solid-phase)	In plates (96,384 wells and beyond) (solution or solid-phase)
Number of compounds	From ca. 330 000 (reasonable) to over 10^{12} (unreasonable)	Limit set by the number of building blocks and the manageability of plates
Biological system	Simple (and highly specific)	Might be complex (including in vivo)
Screening	Highly throughput not necessary	Highly throughput required
Molecular diversity	Via the number of compounds in the mixture	Via the variable structures of building blocks
Special requirement	Synthesis can be manual or robotic	Synthesis automated

the way from the hit, to the lead and finally to the drug candidate requires the classical expertise of the organic chemist including QSAR and molecular modeling for ligand optimisation. Parallel synthesis of analogues is a convenient way to accelerate this process too, especially if it is combined with an efficient analytical throughput such as LC–MS.

Although the major trend is towards the synthesis of large numbers of individual compounds on formats compatible with rapid screening on preliminary ELISA, binding or enzymatic assays, compound mixtures of variable size obtained by truly combinatorial synthesis are still a valuable means to reduce the number of biological tests provided enough investment is made in both safe deconvolution and reliable analytical check.

Acknowledgements

JAB is specially indebted to all the scientists at the Servier Research Institutes that permit the development of the combinatorial chemistry, beside the coauthors of the present review, particularly Drs. Jean-Michel Henlin and Nathalie Kucharczyk, Ms. Christine Desmet and Armelle Loynel and Drs. Philippe Hennig and Florent Meunier.

References

[1] M.R. Pavia, S.P. Hollinshead, H.V. Meyers, S.P. Hall, Chimia 51 (1997) 826.
[2] S. Shi, X.Y. Xiao, A.W. Czarnik, Biotechnol. Bioengineer. 61 (1998) 7.
[3] S.R. Wilson, in: S.R. Wilson (Ed.), Combinatorial Chemistry, John Wiley and Sons, New York, 1997, p. 1.
[4] R.W. Sullivan, C.G. Bigam, P.E. Erdman, S.S. Palanki, D.W. Anderson, M.E. Goldman, L.J. Ransome, M.J. Suto, J. Med. Chem. 41 (1998) 413.
[5] A. Furka, F. Sebestyen, M. Asgedom, D. Dibo, Int. J. Pept. Protein Res. 37 (1991) 487.
[6] R.A. Houghten, C. Pinilla, S.E. Blondelle, J.R. Appel, C.T. Dooley, J.H. Cuervo, Nature 354 (1991) 84.
[7] J.A. Boutin, P. Hennig, P.H. Lambert, S. Bertin, L. Petit, J.P. Mahieu, B. Serkiz, J.P. Volland, J.L. Fauchère, Anal. Biochem. 234 (1996) 126.
[8] W.L. Finch, Ann. Rep. Combin. Chem. Mol. Divers. 1 (1997) 59.
[9] H.M. Geysen, S.J. Rodda, T.J. Mason, Mol. Immunol. 23 (1986) 709.
[10] K.S. Lam, S.E. Salmon, E.M. Hersh, V.J. Hruby, V.M. Kazmierski, R.J. Knapp, Nature 354 (1991) 82.
[11] J.L. Fauchère, J.A. Boutin, J.M. Henlin, N. Kuchartzyk, J.C. Ortuno, Chem. Intell. Lab. System 43 (1998) 43.
[12] P.L. Zhao, R.B. Nachar, J.A. Bolognese, K. Chapman, J. Med. Chem. 39 (1996) 350.
[13] J.A. Boutin, A.L. Fauchère, Trends Pharmacol. Sci. 17 (1996) 8.
[14] G. Barany, R.B. Merrifield, in: E. Gross, J. Meienhofer (Eds.), The Peptides. Analysis, Synthesis, Biology, Academic Press, New York, 1979, pp. 1–284.
[15] C. Pinilla, J.R. Appel, R.A. Houghten, Biochem. J. 301 (1994) 847.
[16] S.E. Salmon, K.S. Lam, M. Lebl, A. Kandola, P.S. Khattri, S. Wade, M. Patek, P. Kocis, V. Krchnak, D. Thorpe, S. Felder, Proc. Natl. Acad. Sci. USA 90 (1993) 11708.
[17] E. Bianchi, A. Folgori, A. Wallace, M. Nicotra, S. Acali, A. Phalipon, G. Barbato, R. Bazzo, R. Cortese, F. Felici, A. Pessi, J. Mol. Biol. 247 (1995) 154.
[18] D. Chargelegue, O.E. Obeid, S.C. Hsu, M.D. Shaw, A.N. Denbury, G. Taylor, M.W. Steward, J. Virol. 72 (1998) 2040.
[19] E. Erb, K.D. Janda, S. Brenner, Proc. Natl. Acad. Sci. USA 91 (1994) 11422.
[20] J.M. Kerr, S.C. Banville, R.N. Zuckermann, Bioorg. Med. Chem. Lett. 3 (1993) 463.

[21] K.S. Lam, M. Lebl, V. Krchnak, S. Wade, F. Abdul-Latif, R. Ferguson, C. Cuzzocrea, K. Wertman, Gene 137 (1993) 13.
[22] W.Y. Wong, H.B. Sheth, A. Holm, R.S. Hodges, R.T. Irvin, Methods 6 (1994) 404.
[23] A. Wallace, S. Altamura, C. Toniatti, A. Vitelli, E. Bianchi, P. Delmastro, G. Ciliberto, A. Pessi, Peptide Res. 7 (1994) 27.
[24] T. Carell, E.A. Wintner, A.J. Sutherland, J. Rebek, Y.M. Dunayevskiy, P. Vouros, Chem. Biol. 2 (1995) 171.
[25] G. Ferry, J.A. Boutin, G. Atassi, J.L. Fauchère, G.C. Tucker, Mol. Div. 2 (1996) 135.
[26] J.D. McBride, N. Freeman, G.J. Domingo, R.J. Leatherbarrow, J. Mol. Biol. 259 (1996) 819.
[27] E.J. Moran, S. Sarshar, J.F. Cargill, M.M. Shahbaz, A. Lio, A.M.M. Mjalli, R.W. Armstrong, J. Am. Chem. Soc. 117 (1995) 10787.
[28] R.A. Owens, P.D. Gesellchen, B.J. Houchins, R.D. DiMarchi, Biochem. Biophys. Res. Commun. 181 (1991) 402.
[29] Z. Songyang, S. Blechner, N. Hoagland, M.F. Hoekstra, H. Piwnica-Worms, L.C. Cantley, Current Biol. 4 (1994) 973.
[30] Z. Songyang, B. Margolis, M. Chaudhuri, S.E. Shoelson, L.C. Cantley, J. Biol. Chem. 270 (1995) 14863.
[31] W. Tegge, R. Frank, F. Hofmann, W.R.G. Dostmann, Biochemistry 34 (1995) 10569.
[32] J.H. Till, R.S. Annan, S.A. Carr, W.T. Miller, J. Biol. Chem. 269 (1994) 7423.
[33] A. Wallace, K.S. Koblan, K. Hamilton, D.J. Marquis-Omer, P.J. Miller, S.D. Mosser, C.A. Omer, M.D. Schaber, R. Cortese, A. Oliff, J.B. Gibbs, A. Pessi, J. Biol. Chem. 271 (1996) 31306.
[34] J. Wu, Q.N. Ma, K.S. Lam, Biochemistry 33 (1994) 14825.
[35] S. Mukhija, L. Germeroth, J. Schneider-Mergener, B. Erni, Eur. J. Biochem. 254 (1998) 433.
[36] R.A. Puras-Lutzke, N.A. Eppens, P.A. Weber, R.A. Houghten, R.H.A. Plasterk, Proc. Natl. Acad. Sci. USA 92 (1995) 11456.
[37] J.A. Boutin, W. Marande, L. Petit, A. Loynel, C. Desmet, E. Canet, J.L. Fauchère, Cell. Signal. 11 (1999) 59.
[38] Y.W. Cheung, C. Abell, S. Balasubramanian, J. Am. Chem. Soc. 119 (1997) 9568.
[39] K.S. Lam, J. Wu, Q. Lou, Int. J. Prot. Peptide Res. 45 (1995) 587.
[40] Q. Lou, M. Leftwich, K.S. Lam, Bioorg. Med. Chem. 5 (1996) 677.
[41] K. Luo, P. Zhou, H.F. Lodish, Proc. Natl. Acad. Sci. USA 92 (1995) 11761.
[42] M. Wigger, J.P. Narocki, C.H. Watson, J.R. Eyker, S.A. Benner, Rapid Commun. Mass Spectrom. 11 (1997) 1749.
[43] M. Renil, M. Ferreras, J.M. Delaisse, N.T. Foged, M. Meldal, J. Peptide Sci. 4 (1998) 195.
[44] B. Déprez, X. Williard, L. Bourel, H. Coste, F. Hyafil, A. Tartar, J. Am. Chem. Soc. 117 (1995) 5405.
[45] C.T. Dooley, R.A. Houghten, Life Sci. 52 (1993) 1509.
[46] C.T. Dooley, N.N. Chung, B.C. Wilkes, P.W. Schiller, J.M. Bidlack, G.W. Pasternak, R.A. Houghten, Science 266 (1994) 2019.
[47] J. Gao, X. Cheng, R. Chen, G.B. Sigal, J.E. Bruce, B.L. Schwartz, S.A. Hofstadler, G.A. Anderson, R.D. Smith, G.M. Whitesides, J. Med. Chem. 39 (1996) 1949.
[48] B. Neustadt, A. Wu, E.M. Smith, T. Nechuta, A. Fawzi, H. Zhang, A.K. Ganguly, Bioorg. Med. Chem. Lett. 5 (1995) 2041.
[49] Z. Songyang, S.E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W.G. Haser, F. King, T. Roberts, S. Ratnofsky, R.J. Lechleider, B.G. Neel, R.B. Birge, J.E. Fajardo, M.M. Chao, H. Hanafusa, B. Schaffhausen, L.C. Cantley, Cell 72 (1993) 767.
[50] Z. Songyang, S.E. Shoelson, J. McGlade, P. Olivier, T. Pawson, X.R. Bustello, M. Barbacid, H. Sabe, H. Hanafusa, T. Yi, R. Ren, D. Baltimore, S. Ratnofsky, R.A. Feldman, L.C. Cantley, Mol. Cell. Biol. 14 (1994) 2777.
[51] Z. Songyang, K.L. Carraway, M.J. Eck, S.C. Harrison, R.A. Feldman, M. Mohammodi, J. Schlessinger, S.R. Hubbard, B.J. Mayer, L.C. Cantley, Nature 373 (1995) 536.
[52] A. Kramer, R. Volkmer-Engert, R. Malin, U. Reineke, J. Schneider-Mergener, Peptide Res. 6 (1993) 314.
[53] R.S. Youngquist, G.R. Fuentes, M.P. Lacey, T. Keough, J. Am. Chem. Soc. 117 (1995) 3900.
[54] T. Tana, N. Kamikawaji, C.J. Savoie, T. Sudo, Y. Kinoshita, T. Sasazuki, J. Hum. Genet. 43 (1998) 14.
[55] K. Udaka, K.H. Wiesmuller, S. Kienle, G. Jung, P. Walden, J. Exp. Med. 181 (1995) 2097.
[56] C.K. Jayawickreme, G.F. Graminski, J.M. Quillan, M.R. Lerner, Proc. Natl. Acad. Sci. USA 91 (1994) 1614.
[57] C.K. Jayawickreme, G.F. Graminski, J.M. Quillan, M.R. Lerner, J. Biol. Chem. 269 (1994) 29846.
[58] J.M. Quillan, C.K. Jayawickreme, M.R. Lerner, Proc. Natl. Acad. Sci. USA 92 (1995) 2894.
[59] G.L. Hortin, W.D. Staatz, S.A. Santoro, Biochem. Int. 26 (1992) 731.
[60] S. Ostergaard, A. Holm, Mol. Div. 3 (1997) 17.
[61] J. Eichler, A.W. Lucka, R.A. Houghten, Peptide Res. 7 (1994) 300.
[62] L. Wilson-Lingardo, P.W. Davis, D.J. Ecker, N. Hebert, O. Acevedo, K. Sprankle, T. Brennan, L. Schwarc, S.M. Freier, J.R. Wyatt, J. Med. Chem. 39 (1996) 2720.
[63] J. Jiracek, A. Yiotakis, B. Vincent, F. Checler, V. Dive, J. Biol. Chem. 271 (1996) 19606.
[64] J.J. Burbaum, M.H.L.J. Ohlmeyer, J.C. Reader, I. Henderson, L.W. Dillard, G. Li, T.L. Randle, N.H. Sigal, D. Chelsky, J.J. Baldwin, Proc. Natl. Acad. Sci. USA 92 (1995) 6027.
[65] K.C. Appell, T.D.Y. Chung, K.J. Solly, D. Chelsky, J. Biomolecular Screening 3 (1998) 19.
[66] S. Chakravarty, B.J. Mavunkel, R.R. Goehring, D.J. Kyle, Immunopharmacol. 33 (1996) 61.
[67] X. Cheng, R. Chen, J.E. Bruce, B.L. Schwartz, G.A. Anderson, S.A. Hofstadler, D.C. Gale, R.D. Smith, J. Gao, G.B. Sigal, M. Mammen, G.M. Whitesides, J. Am. Chem. Soc. 117 (1995) 8859.
[68] V.S. Goodfellow, C.P. Laudeman, J.I. Gerrity, M. Burkard, E. Strobel, J.S. Zuzack, D.A. McLeod, Mol. Div. 2 (1996) 97.
[69] P.W. Smith, J.Y.Q. Lai, A.R. Whittington, B. Cox, J.G. Houston, C.H. Stylli, M.N. Banks, P.R. Tiller, Bioorg. Med. Chem. Lett. 4 (1994) 2821.
[70] M.J. Kurth, L.A. Ahlberg Randall, C. Chen, C. Melander,

R.B. Miller, K. McAlister, G. Reitz, R. Kang, T. Nakatsu, C. Green, J. Org. Chem. 59 (1994) 5862.
[71] R.N. Zuckerman, E.J. Martin, D.C. Spellmeyer, G.B. Stauber, K.R. Shoemaker, J.M. Kerr, G.M. Figliozzi, D.A. Goff, M.A. Siani, R.J. Simon, S.C. Banville, E.G. Brown, L. Wang, L.S. Richter, W.H. Moos, J. Med.Chem. 37 (1994) 2678.
[72] J.L. Silen, A.T. Lu, D.W. Solas, M.A. Gore, D. Maclean, N.H. Shah, J.M. Coffin, N.S. Bhinderwala, Y. Wang, K.T. Tstutsui, G.C. Look, D.A. Campbell, R.L. Hale, M. Navre, C.R. Luca-Flaherty, Antimicrobial Agents Chemother. 42 (1998) 1447.
[73] H. An, T. Wang, V. Mohan, R.H. Giffrey, P.D. Cook, Tetrahedron 54 (1998) 3999.
[74] M. Stankova, M. Lebl, Mol. Div. 2 (1996) 75.
[75] S.M. Dankwardt, T.M. Phan, J.L. Krstenansky, Mol. Div. 1 (1995) 113.
[76] J.A. Loo, Eur. Mass Spectrom. 3 (1997) 93.
[77] B.J. Egner, M. Bradley, Drug Discovery Today 2 (1997) 102.
[78] P.C. Andrews, J. Boyd, R. Ogorzalek Loo, R. Zhao, C.Q. Zhu, K. Grant, S. Williams, in: V.J.W. Crabb (Ed.), Tech. Protein Chem, Academic Press, San Diego, CA., USA, 1994, pp. 485–492.
[79] J.A. Boutin, I. Gesson, J.M. Henlin, S. Bertin, P.H. Lambert, J.P. Volland, J.L. Fauchère, Mol. Div. 3 (1997) 43.
[80] Y.M. Dunayevskiy, P. Vouros, E.A. Wintner, G.W. Shipps, T. Carell, J. Rebek, Proc. Natl. Acad. Sci. USA 93 (1996) 6152.
[81] Y. Dunayevskiy, P. Vouros, T. Carell, E.A. Winter, J. Rebek, Anal. Chem. 67 (1995) 2906.
[82] S.C. Pomerantz, J.A. McCloskey, T.M. Tasarow, B.E. Eaton, J. Am. Chem. Soc. 119 (1997) 3861.
[83] W.J. Haap, J.W. Metzger, C. Kempter, G. Jung, Mol. Div. 3 (1997) 29.
[84] R. Fahti, R. Patel, A.F. Cook, Mol. Div. 2 (1996) 125.
[85] P.-H. Lambert, J.A. Boutin, S. Bertin, J.-L. Fauchère, J.-P. Volland, Rapid Commun. Mass Spectrom. 11 (1997) 1971.
[86] J.W. Metzger, K.H. Wiesmüller, V. Gnau, J. Brünjes, G. Jung, Angew. Chem. Int. Ed. Engl. 32 (1993) 894.
[87] J.W. Metzger, C. Kempter, K.H. Wiesmüller, G. Jung, Anal. Biochem. 219 (1994) 261.
[88] J.P. Nawrocki, M. Wigger, C.H. Watson, T.W. Hayes, M.W. Senko, S.A. Benner, J.R. Eyler, Rapid Commun. Mass Spectrom. 10 (1996) 1860.
[89] S. Stevanovic, K.H. Wiesmüller, J. Metzger, A.G. Beck-Sickinger, G. Jung, Bioorg. Med. Chem. Lett. 3 (1993) 431.
[90] B.E. Winger, J.E. Campana, Rapid Commun. Mass Spectrom. 10 (1996) 1811.
[91] R.N. Zuckermann, J.M. Kerr, M.A. Siani, S.C. Bandeville, D.V. Santi, Proc. Natl. Acad. Sci. USA 89 (1996) 126.
[92] P.A. Demirev, R.A. Zubarev, Anal. Chem. 69 (1997) 2893.
[93] K.F. Blom, Anal. Chem. 69 (1997) 4354.
[94] R. Volmer-Engert, B. Ehrhard, J. Hellwig, A. Kramer, W. Höhne, J. Schneider-Mergener, Lett. Peptide Sci. 1 (1994) 243.
[95] A. Quesnel, A. Delmas, Y. Trudelle, Anal. Biochem. 231 (1995) 182.
[96] T. Carell, E.A. Wintner, A. Bashir-Hashemi, J. Rebek, Angew. Chem. Int. Ed. Engl. 33 (1994) 2059.
[97] B. Yan, J.B. Fell, G. Kumaravel, J. Org. Chem. 61 (1996) 7467.
[98] R.C. Anderson, J.P. Stokes, M.J. Shapiro, Tetrahedron Lett. 30 (1995) 5311.
[99] M.R. Carrasco, M.C. Fitzgerald, Y. Oda, S.B.H. Kent, Tetrahedron Lett. 38 (1997) 6331.
[100] V. Krchnak, M. Lebl, Mol. Div. 1 (1995) 193.
[101] E.R. Felder, D. Poppinger, Adv. Drug Res. 30 (1997) 111.
[102] S.M. Freier, D.A.M. Konings, J.R. Wyatt, D.J. Ecker, J. Med. Chem. 38 (1995) 344.
[103] G. Ferry, J.A. Boutin, P. Hennig, A. Genton, C. Desmet, J.L. Fauchère, G. Atassi, G.C. Tucker, Eur. J. Pharmacol. 351 (1998) 225.
[104] C. Pinilla, J.R. Appel, P. Blanc, R.A. Houghten, BioTechniques 13 (1992) 901.
[105] A. Furka, Drug Dev. Res. 33 (1994) 90.
[106] M. Lebl, V. Krchnak, N.F. Sepetov, B. Seligman, P. Strop, S. Felder, Biopolymers 37 (1995) 177.
[107] K.S. Lam, M. Lebl, V. Krchnak, Chem. Rev. 97 (1997) 411.
[108] R. Frank, Tetrahedron 48 (1992) 9217.
[109] A. Kramer, J. Schneider-Mergener, Meth. Mol. Biol. 87 (1997) 25.
[110] U. Reineke, R. Sabat, A. Kramer, R.D. Stigler, M. Seifert, T. Michel, H.D. Volk, J. Schneider-Mergener, Mol. Div. 1 (1995) 141.
[111] M. Yonemoto, T. Satoh, H. Arawaka, I. Suzuki-Takahashi, Y. Monden, T. Kodera, K. Tanaka, T. Aoyama, Y. Iwasawa, T. Kamei, S. Nishimura, K. Tomimoto, Mol. Pharmacol. 54 (1998) 1.
[112] M.G. Eason, H. Kurose, B.D. Holt, J.R. Raymond, S.B. Liggett, J. Biol. Chem. 267 (1992) 15795.
[113] S. Virolainen, C.C. Jansson, M. Scheinin, Eur. J. Pharmacol. 338 (1997) 293.
[114] J.A. Boutin, J. Thomassin, G. Siest, A. Cartier, Biochem. Pharmacol. 34 (1985) 2235.
[115] J.A. Boutin, G. Ferry, A.P. Ernould, P. Maes, G. Remond, M. Vincent, Eur. J. Biochem. 214 (1993) 853.
[116] A.P. Ernould, G. Ferry, J.M. Barret, A. Genton, J.A. Boutin, Int. J. Peptide Prot. Res. 43 (1994) 496.
[117] J.A. Boutin, W. Marande, M. Goussard, A. Loynel, E. Canet, J.L. Fauchère, Arch. Biochem. Biophys. 354 (1998) 83.
[118] R. Roskoski, P. Ritchie, Arch. Biochem. Biophys. 356 (1998) 167.
[119] J.A. Boutin, Int. J. Biochem. 26 (1994) 1203.
[120] J.A. Boutin, Cell. Signal. 9 (1997) 15.
[121] J.A. Boutin, B. Antoine, G. Siest, J. Pharm. Sci. 83 (1994) 591.
[122] D.M. Bickett, M.D. Green, J. Berman, M. Dezube, A.S. Howe, P.J. Brown, J.T. Roth, G.M. McGeehan, Anal. Biochem. 212 (1993) 58.
[123] R.E. Dolle, Mol. Div. 2 (1996) 223.
[124] T.C. Chou, P. Talalay, Eur. J. Biochem. 115 (1981) 207.
[125] T.C. Chou, P. Talalay, Adv. Enzyme Regul. 22 (1984) 27.
[126] N.S. Gray, L. Wodicka, A.M.W.H. Thunnissen, T.C. Nor-

man, S. Kwon, F.H. Espinoza, D.O. Morgan, G. Barnes, S. LeClerc, L. Meijer, S.H. Kim, D.J. Lockhart, P.G. Schultz, Nature 281 (1998) 533.
[127] J.M. Barret, A.P. Ernould, G. Ferry, A. Genton, J.A. Boutin, Biochem. Pharmacol. 46 (1993) 439.
[128] M. Meldal, I. Svendsen, J. Chem. Soc., (1995) 1591
[129] S. Sasaki, M. Takagi, Y. Tanaka, M. Maeda, Tetrahedron Lett. 37 (1996) 85.
[130] M.A. Sills, Drug Discovery Today 3 (1998) 304.
[131] L.C. Cantley, K.R. Auger, C. Carpenter, B. Duckworth, A. Graziani, R. Kappeler, S. Soltoff, Cell 64 (1991) 281.
[132] M. Sudol, TiBS 21 (1996) 161.
[133] K.L. Lan, N.A. Sarvazyan, R. Taussig, R.G. Mackenzie, P.R. DiBello, H.G. Dohlman, R.R. Neubig, J. Biol. Chem. 273 (1998) 12794.
[134] E. Kostenis, F.Y. Zeng, J. Wess, J. Biol. Chem. 273 (1998) 17886.
[135] A. Weiss, J. Schlessinger, Cell 94 (1998) 277.
[136] T.E. Hebert, S. Moffett, J.P. Morello, T.P. Loisel, D.G. Bichet, C. Barret, M. Bouvier, J. Biol. Chem. 271 (1996) 16384.
[137] D. Dong, N. Noy, Biochemistry 37 (1998) 10691.
[138] S.E. Egan, R.A. Weinberg, Nature 365 (1993) 781.
[139] H.M. Daiman, R.R. Neubig, J. Biol. Chem. 266 (1991) 11025.
[140] J.B. Thompson, S.M. Wade, J.K. Harrison, M.N. Salafranca, R.R. Neubig, J. Pharmacol. Exp. Therap. 285 (1998) 216.
[141] R.P. Loudon, Drug News Perspect. 11 (1998) 73.
[142] S.E. Salmon, R.H. Liu-Stevens, Y. Zhao, M. Lebl, V. Krchnark, K. Wertman, N. Sepetov, K.S. Lam, Mol. Div. 2 (1996) 57.
[143] G.F. Graminski, C.K. Jayawickreme, M.N. Potenza, M.R. Lerner, J. Biol. Chem. 268 (1993) 5957.
[144] R. Obst, C. Münz, S. Stevanovic, H.G. Rammensee, Eur. J. Immunol. 28 (1998) 2432.
[145] D.C. Horwell, W. Howson, G.S. Ratcliffe, D. Rees, Bioorg. Med. Chem. Lett. 3 (1993) 799.
[146] H.M. Geysen, S.J. Rodda, T.J. Mason, G. Tribbick, P.G. Schoofs, J. Immunol. Meth. 102 (1987) 259.
[147] E.J. Ariëns, in: V. Claassen (Ed.), Trends in Drug Research, Pharmacochemistry Library, Elsevier, Amsterdam, 1990, pp. 1–10.
[148] J.L. Fauchère, M. Charton, L.B. Kier, A. Verloop, V. Pliska, Int. J. Peptide Prot. Res. 32 (1988) 269.
[149] M. Sandberg, L. Eriksson, J. Jonsson, M. Sjöström, S. Wold, J. Med. Chem. 41 (1998) 2481.
[150] A.F. Spatola, P. Romanovskis, in: G. Jung (Ed.), Combinatorial Peptide and Non-peptide Libraries, VCH, Weinheim, Germany, 1996, pp. 327–347.
[151] G. Grassy, A. Yasri, B. Calas, A.M. Armbruster, R. Lahana, J.L. Fauchère, in: H. van de Waterbeemd, B. Testa, G. Folkers (Eds.), Computer-assisted Lead Finding and Optimization, Helvetica Chimica Acta, Basel, 1997, pp. 211–221.
[152] C.J. Blankley, Pharm. Pharmacol. Commun. 4 (1998) 139.
[153] Y.C. Martin, M.G. Bures, R.D. Brown, Pharm. Pharmacol. Commun. 4 (1998) 147.
[154] W. Zheng, S.J. Cho, A. Tropsha, J. Chem. Inf. Comput. Sci. 38 (1998) 251.
[155] S.J. Cho, W. Zheng, A. Tropsha, J. Chem. Inf. Comput. Sci. 38 (1998) 259.
[156] H.J. Chin, B. Fell, M.J. Shapiro, J. Tomesch, J.R. Wareing, A.M. Bray, J. Org. Chem. 62 (1997) 538.
[157] H.C. Dorn, Anal. Chem. 56 (1984) 747A.
[158] J. Chin, J.B. Fell, M. Jarosinski, M.J. Shapiro, J.R. Wareing, J. Org. Chem. 63 (1998) 386.
[159] A.W. Czarnik, Anal. Chem. 70 (1998) 378A.
[160] P.L. Myers, Curr. Opin. Biotechnol. 8 (1998) 701.
[161] E. Apletalina, J. Appel, N.S. Lamango, R.A. Houghten, I. Lindberg, J. Biol. Chem. 273 (1998) 26589.